Thesis / Reports Loftgaarden, D.O.

THE RELATIONSHIP BETWEEN THE NATIONAL FIRE DANGER RATING SYSTEM INDEXES & FIRE OCCURRENCE - AN EXPLORATORY STUDY

COOPERATIVE AGREEMENT WITH SYSTEMS FOR ENVIRONMENTAL MANAGEMENT
No. INT=87228 - SEM
Co-op Contact: Dr. Donald Loftsgaarden
FS Contact: Richard C. Rothermel

United States Department of Agriculture | Forest Service | INT

1580 Grants and Agreements

Date: August 3, 1987

Subject: Cooperative Agreement No. INT-87228, Final Report with Systems for Environmental Management

To: Station Director
INT - Ogden
THRU: Dr. R. G. ~~Krebill~~, ASD-N
RK

TO: ~~B&EA 9/3/87~~
~~ASD-Admin~~
~~DD~~ seen
Westfornet (for File)

Enclosed is the final report for coop agreement No. INT-87228 with SEM. The principal investigator, Dr. Donald Loftsgaarden, did excellent work in showing relationships between the National Fire Danger Rating System and the probability of fire occurrence on the Lolo National Forest. This coop may be considered complete. We anticipate further work on this problem, examining data from another area, e.g., Mississippi, and the development of a generalized model.

Co-op Contact

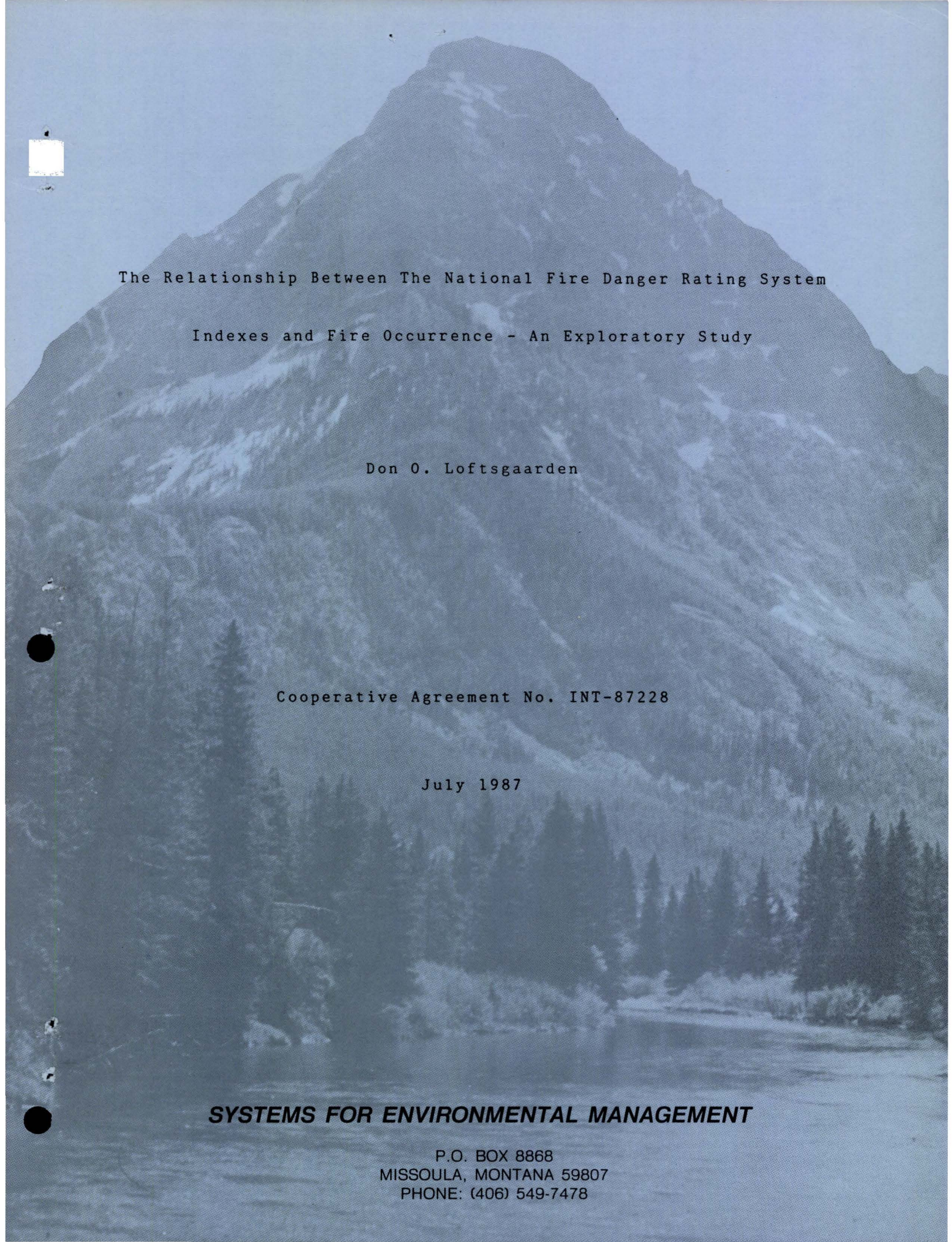
R C Rothermel

FS-Contact

RICHARD C. ROTHERMEL
Project Leader

Enclosure

RECEIVED
AUG 5 - 1987
RESEARCH STATION



The Relationship Between The National Fire Danger Rating System Indexes and Fire Occurrence - An Exploratory Study

Don O. Loftsgaarden

Cooperative Agreement No. INT-87228

July 1987

SYSTEMS FOR ENVIRONMENTAL MANAGEMENT

P.O. BOX 8868
MISSOULA, MONTANA 59807
PHONE: (406) 549-7478

SYSTEMS FOR ENVIRONMENTAL MANAGEMENT

A Natural Resources Research Group

Systems for Environmental Management (SEM) is a non-profit research and educational corporation based in Missoula, Montana, a regional center for natural resources agencies. Founded in 1977, SEM works cooperatively and under contract with the U.S. Forest Service, National Park Service, Bureau of Land Management, and Bureau of Indian Affairs, as well as state and academic institutions, private individuals and organizations.

SEM's diverse professional staff offers a wide range of research capabilities within the natural resources management field. The full-time staff includes specialists in fire behavior, fire history, recreation management, geography, plant ecology, forestry, meteorology, data analysis, and computer science. SEM also maintains a pool of professional affiliates that can be utilized for projects requiring additional expertise. Areas of recent research emphasis include wilderness recreation management, fire history and ecology, fire planning, and development of computerized resource management tools.

The Relationship Between The National Fire Danger Rating System Indexes and Fire Occurrence - An Exploratory Study

Don O. Loftsgaarden

Cooperative Agreement No. INT-87228

July 1987

This report is submitted in fulfillment of section III (paragraph B) of cooperative aid agreement No. INT-87228 between the Intermountain Research Station and Systems for Environmental Management, Missoula, Montana.

This study was exploratory in nature and this report, for the most part, simply summarizes results of this exploration. It is an in-house report, not a technical report.

Table of Contents

	Page No.
Introduction	2
The Data	3
Weather Data File--description and summary statistics	3
Fire Data File--description and summary statistics	4
Merged Data File--description and summary statistics	5
Summary Statistics for Missoula District and Missoula Weather Station	7
Logistic Models (General Discussion)	9
Poisson Models (General Discussion)	11
Logistic Models for each District using ERC	13
Models based on each District's own ERC values	15
Models based on ERC's from the Missoula Station	15
Tables giving observed and estimated results for models based on each District's own ERC's	16
Tables giving observed and estimated results for models based on the Missoula Station's ERC's	17
Examples of Poisson Models (Superior District)	18
Summary Statistics for Lolo National Forest Days	20
Logistic Models for Lolo National Forest Days	21
Tables giving observed and estimated results for these logistic models	22
A Comparison of Three Logistic Models for the Probability of a Fire Day in the Lolo National Forest	24
Examples of Poisson Models (Lolo National Forest Days)	26
Logistic Models and Results for Large Fire Days-- Lolo National Forest	27
Scores for Evaluating Predictions from Logistic Models	31
Scores for the Logistic Models Given in This Report	32
References	34
Appendix A--List of Variables Taken from the Original Data File	35
Appendix B--List of Programs and Data Files	36
Appendix C--Items Worthy of Further Study	42

INTRODUCTION

The overall objective of this exploratory study was to examine the relationship between National Fire Danger Rating System (NFDRS) indexes and fire occurrence and size. Lolo National Forest was chosen to study. Historical data for years 1970-1985 were taken from data bases at the Fort Collins Computer Center (FCCC): the National Fire Weather Data Library (NFWDL) and the National Fire Occurrence Data Library (NFODL).

In phase one, logistic models for predicting the probability of at least one fire on a day as a function of one or more of the fire indexes were developed at both the District level and the Forest level. The purpose was to see if such models could successfully model the probability of a fire day. Logistic models have been used successfully by Martell [1] for this purpose.

In phase two, a quadratic scoring rule was used as a measure of how well these models work and for comparing different models. Each day a logistic model gives a prediction of the probability of a fire that day. The day either has a fire or it doesn't have a fire. This is exactly the problem that meteorologists have in predicting precipitation each day. The quadratic scoring rule was developed in that setting by Winkler and Murphy [2].

In phase 3, Poisson probability models were developed for the number of fires each day, i.e., probability of 0 fires, probability of 1 fire, etc. The parameter for the Poisson was estimated using the estimated probability of a fire day from the logistic model. Martell [1] has used this model successfully.

The results of this exploratory study indicate that the approach outlined above looks very promising and seems to warrant continued

study. A reasonable second step would be to examine another National Forest. Other directions for future study are outlined in Appendix C.

THE DATA

Lolo National Forest has six districts: Missoula (03), Ninemile (04), Plains (05), Seeley Lake (06), Superior (07), and Thompson Falls (08). There are five weather stations: Missoula (241507), Plains (241206), Superior (241302), Seeley Lake (241508), and Thompson Falls (241209). The Ninemile District uses weather from the Missoula station.

Weather and fire information for 16 years (1970-1985) was used in this study.

WEATHER DATA FILE

Weather information, including NFDRS indexes for each day during the fire season at each station, was obtained from the National Fire Weather Data Library. Each year the data available at each station was from the time the weather station was operating in the Spring until some time in October. The number of days of information available each year varied from station to station, as well as from year to year. Occasionally data were missing from a day or days in the middle of the fire season. This missing information was filled in, using simple interpolation. A complete list of variables taken from the NFWDL is given in Appendix A.

The number of days of weather information at each station for the 16 years follows:

Station	241206	241209	241302	241507	241508	
Number of days of weather	2458	1556	2380	2457	1622	10477

For all stations combined the breakdown of number of days of weather information by year and by month follows:

Year	70	71	72	73	74	75	76	77	78	79
Number of days of weather	714	692	608	512	552	552	576	786	753	784

80	81	82	83	84	85	
775	797	651	665	485	575	10477

Month	May	June	July	August	Sept.	Oct.	
Number of days of weather	1172	1759	2449	2449	2059	589	10477

FIRE DATA FILE

Data for fires were obtained from the National Fire Occurrence Data Library. This file had one case for each fire in each district for years 1970-1985, inclusive. A case was added to this file for each day in each district that did not have a fire. Each day in each district then had at least one case. A list of variables taken from the NFODL is given in Appendix A.

At this point this data file had 18,614 cases of which 16,172 were for a day with no fires and 2,442 cases were present where each represented one fire. Some of these cases were eventually not usable because they were for days when no weather and fire indexes were available. The breakdowns of these cases by district, year, month, size, and cause follow:

Breakdowns of the Fire Data File

By District

District	Cases
Missoula	3,098
Ninemile	3,158
Plains	3,033
Seeley Lake	3,104
Superior	3,191
Thompson Falls	3,027
	18,614

By Month

Month	Cases
May	2,989
June	2,936
July	3,222
August	3,529
September	2,939
October	2,999
	18,614

By year

Year	Cases
70	1,199
71	1,134
72	1,160
73	1,297
74	1,179
75	1,120
76	1,149
77	1,150
78	1,126
79	1,229
80	1,125
81	1,198
82	1,115
83	1,114
84	1,174
85	1,145
	18,614

By size

Size	Cases
No fire	16,172
0-.25 acres (A)	1,891
.26-9 acres (B)	465
10-99 acres (C)	51
100-acres (D+)	35
	18,614

By cause

Cause		Cases
Lightning		1,650
Human		792
-equip. use	88	
-smoking	70	
-camp fire	158	
-debris burning	134	
-railroad	154	
-incendiary	29	
-children	58	
-misc.	101	
		2,442

MERGED DATA FILE

The weather data file and fire data file were combined and condensed in two steps. For this study, both weather and fire information must be available in order for a day to be usable.

In step 1, year, month, day, and station number are used to attach the correct weather and fire indexes to each case in the fire file. At this point, each district has several cases in the file for many days.

Much of the information in cases in the same district on the same day is redundant. Also, for the analyses carried out in this study, individual fires are not needed. What is needed is the number of fires and number of large fires (10 acres or larger) each day.

In step 2, cases for the same day in a district are aggregated. In addition, cases with weather and indexes missing are deleted. The number of cases in this merged/aggregated data file was 12,810.

The tables that follow present a variety of information gleaned from this merged data file.

District	Days	Fire days*	Large fire days*	Range of ERC values***
Missoula	2,442	294	13	1 - 60
Ninemile	2,442	294	10	1 - 60
Plains	2,444	180	7	1 - 84
Seeley Lake	1,607	177	6	1 - 54
Superior	2,328	301	21	1 - 53
Thompson Falls	1,547	125	8	1 - 75
	12,810			

* Fire day = a day with at least 1 fire
** Large fire day = a day with at least 1 fire of 10 acres or larger
*** ERC = Energy release component

	Number of Fires								
District	0	1	2	3	4	5	6	7	$\geq$ 8
Missoula	2,148*	227	36	13	6	2	3	3	4
Ninemile	2,148	213	30	21	11	10	2	2	5
Plains	2,264	134	24	13	4	3	1	0	1
Seeley Lake	1,430	113	34	15	5	3	2	2	3
Superior	2,027	202	52	17	11	6	4	2	7
Thompson Falls	1,422	96	16	5	1	4	0	1	2

*The numbers in the table are days

For many analyses in this study, Lolo National Forest is treated as an entity and districts are ignored. When this is done we speak of Lolo National Forest days and the weather and indexes for these days are from the Missoula weather station. For this reason, summary statistics are given for several variables at this station.

Summary Statistics for the Missoula Weather Station (241507)

ERC	Days
0-8	78
9-16	215
17-24	418
25-32	706
33-36	339
37-40	276
41-44	176
45-48	117
49-52	75
53-56	28
57-60	14
	2,442

BI	Days
0-8	132
9-16	97
17-24	342
25-28	346
29-32	443
33-36	286
37-40	241
41-44	164
45-48	92
49-52	93
53-60	105
61-70	57
71-80	22
81-90	17
91-100	4

RHINDEX	Days
0-10	20
11-20	87
21-30	111
31-36	92
37-42	154
43-48	150
49-52	134
53-56	163
57-60	172
61-64	225
65-68	235
69-72	276
73-76	248
77-80	169
81-90	200
91-100	6

SC	Days
0-2	168
3	259
4	439
5	430
6	289
7	204
8	119
9-12	305
13-20	187
21-40	37
41-60	0
101-124	2

IC	Days
0	165
1-4	209
5-8	279
9-12	299
13-16	287
17-20	278
21-24	235
25-28	205
29-32	114
33-36	97
37-40	77
41-50	93
51-60	53
61-70	19
71-80	16
81-90	13
91-100	3

ERC = Energy Release Component
IC = Ignition Component
BI = Burning Index
SC = Spread component
RHINDEX = 100 - Relative Humidity

Summary Statistics for the Missoula Weather Station (241507)

and Missoula District (03)

ERC	Days	Fire days	Large fire days
0-8	78	2 (2.6%)	0 (0%)
9-16	215	9 (4.2%)	0
17-24	418	17 (4.1%)	0
25-32	706	58 (8.2%)	2 (.3%)
33-36	339	52 (15.3%)	1 (.3%)
37-40	276	45 (16.3%)	1 (.4%)
41-44	176	48 (27.3%)	3 (1.7%)
45-48	117	25 (21.4%)	2 (1.7%)
49-52	75	23 (30.7%)	3 (4.0%)
53-56	28	10 (35.7%)	0
57-60	14	5 (35.7%)	1 (7.1%)
	2,442	294 (12.0%)	13 (.5%)

Number of Fires

ERC	0	1	2	3	4	5	6	7	≥8	Total
0-8	76	2								78
	.974	.026								
9-16	206	9								215
	.958	.042								
17-24	401	17								418
	.959	.041								
25-32	648	51	5	0	2					706
	.918	.072	.007	.000	.003					
33-36	287	41	5	3	0	1	0	0	2	339
	.847	.121	.015	.009	.000	.003	.000	.000	.006	
37-40	231	31	8	4	1	0	1			276
	.837	.112	.029	.014	.004	.000	.004			
41-44	128	36	9	0	2	0	0	1		176
	.727	.205	.051	.000	.011	.000	.000	.006		
45-48	92	15	6	3	0	0	1			117
	.786	.128	.051	.026	.000	.000	.009			
49-52	52	15	1	2	1	1	0	2	1	75
	.693	.200	.013	.027	.013	.013	.000	.027	.013	
53-56	18	6	2	1	0	0	0	0	1	28
	.643	.214	.071	.036	.000	.000	.000	.000	.036	
57-60	9	4	0	0	0	0	1			14
	.643	.286	.000	.000	.000	.000	.071			
	2,148	227	36	13	6	2	3	3	4	2,442
	.880	.093	.015	.005	.002	.001	.001	.001	.002	

LOGISTIC MODELS (General Discussion)

One main objective of this study was to investigate the use of logistic models to obtain estimates for the probability of a fire day as a function of one or more NFDRS indexes. This was done at the district level as well as at the Forest level. A very brief general discussion of logistic models in this setting is given here.

For each day we consider an indicator random variable Y where Y is 0 if there is no fire and 1 if there is one or more fires. The probability function for Y is

y	0	1
P(Y=y)	1-P	P

where P = the probability of a fire that day.

Note that $E(Y) = P$. Often one assumes that $E(Y) = \alpha + \beta \cdot x$. Here we assume that

$$P = E(Y) = \frac{e^{\beta_0 + \beta_1 \cdot x}}{1 + e^{\beta_0 + \beta_1 \cdot x}} ,$$

where x is an NFDRS index.

When P is plotted against x, one obtains the well known logistic curve. It is bounded below and above by 0 and 1 which is appropriate since P is a probability. A specific example of such a curve may be found a few pages ahead in this report. We note in passing that

$$P' = \ln \left(\frac{P}{1-P} \right) = \beta_0 + \beta_1 \cdot x$$

This is called the logit transformation and it is often used to check on the validity of the logistic model geometrically.

The data for n days are given by $(x_1, y_1), (x_2, y_2), \ldots (x_n, y_n)$ where x_i = index value on day i and y_i is 0 or 1 depending on whether there is a fire that day or not. These raw observations can be used to obtain maximum likelihood estimators b_0 and b_1, for β_0 and β_1. A second method of using these data is to group the observations using non-overlapping intervals for the index variable. One then uses the number observations and number of fire days within each interval as the raw data. The procedure PROBIT in SPSSX was used to get the maximum likelihood estimates b_0 and b_1 for β_0 and β_1. Some models were constructed using the original observations and some were constructed using grouped data. Some models were constructed using each method. There was very little difference in the models obtained using either method.

The logistic model, $\hat{P}$, for P has the following form:

$$\hat{P} = \frac{e^{b_0 + b_1 x}}{1 + e^{b_0 + b_1 x}} .$$

$\hat{P}$ is the estimated probability of a fire day as a function of x.

Logistic models with two or more index variables in them are easily constructed. With two index variables, the form of the logistic equation would be

$$\hat{P} = \frac{e^{b_0 + b_1 x_1 + b_2 x_2}}{1 + e^{b_0 + b_1 x_1 + b_2 x_2}} .$$

Example

As an example, we consider a logistic model developed for Lolo National Forest Days using the index ERC. The model is

$$\hat{P} = e^{-3.4877 + .09110 \cdot ERC} \Big/ (1 + e^{-3.4877 + .09110 \cdot ERC})$$

The following table gives values of $\hat{P}$ for selected values of ERC and the figure contains a plot of $\hat{P}$ versus ERC.

ERC	$\hat{P}$	ERC	$\hat{P}$	ERC	$\hat{P}$
0	.030	44	.627	84	.985
8	.060	56	.834	92	.993
20	.159	68	.937	100	.996
32	.361	76	.969		

POISSON MODELS (General Discussion)

For a given day and given index variable, let Z be the number of fires that day. For this study, we assumed a Poisson model for Z, i.e.,

$$P(Z = z) = \frac{\lambda^z e^{-\lambda}}{z!} \qquad z = 0, 1, 2, \ldots$$

Note that z = 0 occurs if and only if there is no fire on the day being considered. Thus we have

$$P(Z = 0) = e^{-\lambda} = 1-P$$

from which we find that $\lambda = -\ln(1-P)$.

Using this last fact, we obtain the following estimate for λ:

$$\hat{\lambda} = -\ln(1-\hat{p}).$$

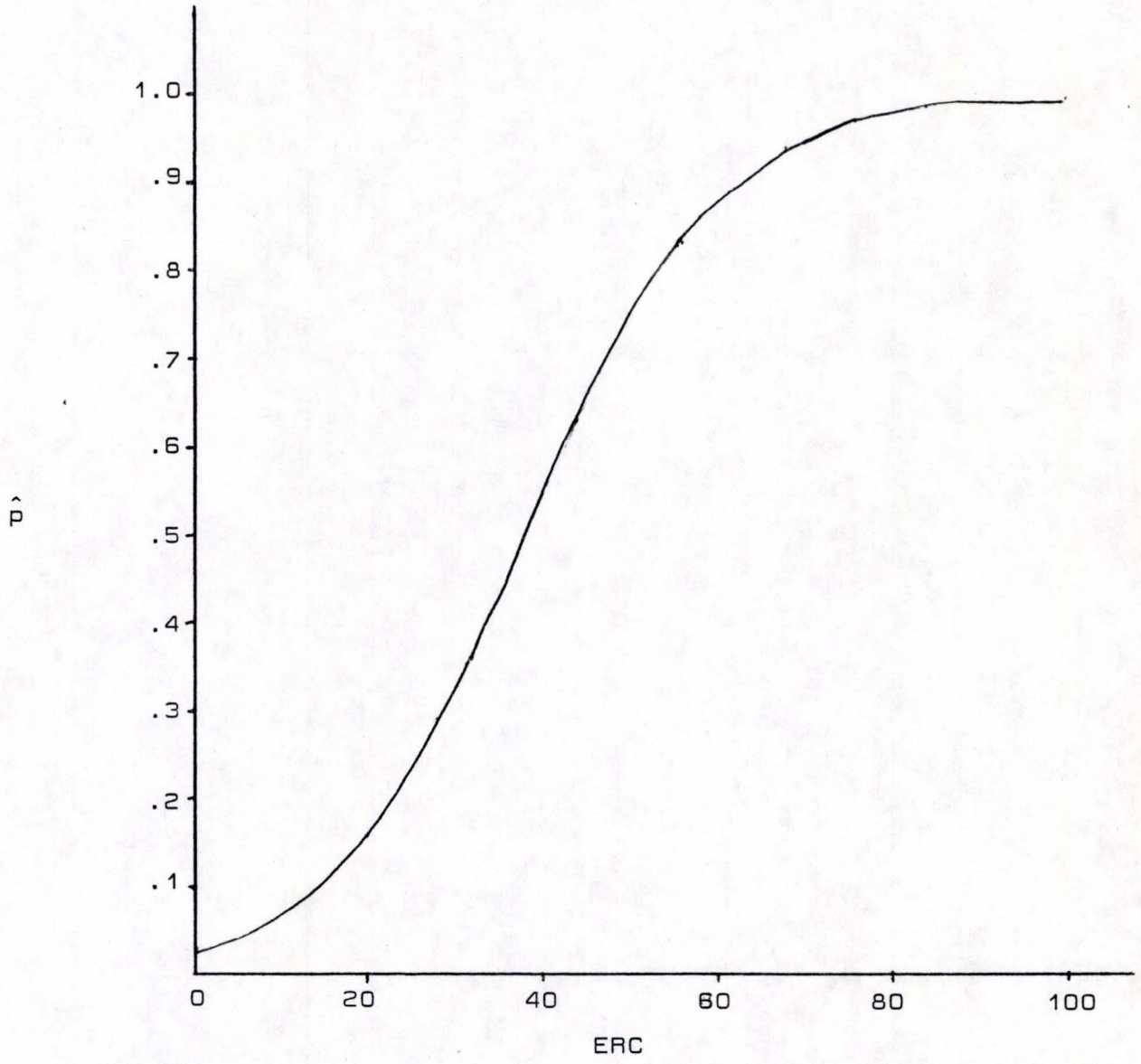
Once we have λ the Poisson model is completely specified and estimates of the probability of z fires (z = 0, 1, 2, . . .) are easily obtained.

For ERC = 44, $\hat{P} = .627$, $\hat{\lambda} = .986$ and

$$P(Z=z) = \frac{.986^z e^{-.986}}{z!} \qquad z = 0, 1, 2, 3, \ldots$$

Figure 1. Lolo National Forest Days

$$\hat{p} = \frac{e^{-3.4877 + .0911ERC}}{1 + e^{-3.4877 + .0911ERC}}$$



This model gives the following probability table.

Z	0	1	2	3	4	≥ 5
P(Z=z)	.373	.368	.181	.060	.015	.003

LOGISTIC MODELS FOR EACH DISTRICT USING ERC

Much of the emphasis of this study was on building models for Lolo Forest days as opposed to models for individual districts. However, in this section two sets of models for individual districts are presented. These models were constructed using only the index variable ERC which was found to be the best index variable to use in these logistic models. This point will be discussed further in the sections on models for Lolo Forest days. No district models were developed for the probability of a large fire day because of the small number of large fires in each district.

The first set of models was developed with each district's model based on the ERC values from its own weather station. Recall that the Ninemile district used the Missoula station for its weather/index values.

The second set of models was developed using the ERC values from the Missoula station in all districts. The Missoula and Ninemile models are, of course, the same. These second models were all based on 2,442 days--the number of days of weather/index values available at the Missoula station. The number of days of weather/index values available at a couple of stations for the first set of models was several hundred less than this.

The following table gives the correlations between ERC values at the five weather stations. The second number in each cell is the number of observations used in calculating the correlation coefficient.

CORRELATION OF ERC'S AT DIFFERENT WEATHER STATIONS

	Missoula	Plains	Seeley Lake	Superior
Plains	.8368 (2349)			
Seeley Lake	.7918 (1602)	.7387 (1605)		
Superior	.8553 (2302)	.8216 (2291)	.7628 (1577)	
Thompson Falls	.7122 (1545)	.7659 (1547)	.7821 (1407)	.7505 (1531)

In all districts, the second model was at least as good and, in a couple of cases, substantially better than the first model constructed. The comparison of models was made using a quadratic scoring rule which is discussed later in this report.

Each model has the form

$$\text{Probability of a fire day} = \hat{P} = \frac{e^{c + d \cdot ERC}}{1 + e^{c + d \cdot ERC}} \quad .$$

MODELS BASED ON EACH DISTRICT'S OWN ERC VALUES

District	c	d
Missoula	-4.27604	.07034
Ninemile	-4.56344	.07844
Plains	-4.54216	.04774
Seeley Lake	-3.84100	.05874
Superior	-5.05698	.10120
Thompson Falls	-4.57070	.05640

MODELS BASED ON ERC VALUES FROM THE MISSOULA STATION

District	c	d
Missoula	-4.27604	.07034
Ninemile	-4.56344	.07844
Plains	-4.72406	.06742
Seeley Lake	-4.85526	.07370
Superior	-4.71366	.08466
Thompson Falls	-5.51708	.08090

The data supplied to the program PROBIT in SPSSX to obtain the models were summarized by intervals of values for each index as shown in the tables immediately following. It appears that a particular choice of intervals has very little effect on the model produced. Later in this report, a model is constructed without grouping the data at all and it is nearly identical to the model constructed with grouped data.

In the tables on the next several pages, the observed fraction of fire days is compared with the estimated fraction of fire days. Also, the observed number of fire days is compared with the estimated number of fire days. Estimates are obtained using the logistic models.

MODELS BASED ON EACH DISTRICT'S ERC

		- -Missoula District- - (Missoula ERC's) Frac. fire days (days)		- Ninemile District - (Missoula ERC's) Frac. fire days (days)	
ERC	Days	Observed	Estimated	Observed	Estimated
0-8	78	.026 (2)*	.018 (1.4)*	.013 (1)	.014 (1.1)
9-16	215	.042 (9)	.031 (6.7)	.014 (3)	.026 (5.6)
17-24	418	.041 (17)	.054 (22.4)	.041 (17)	.048 (19.9)
25-32	706	.082 (58)	.091 (63.9)	.092 (65)	.086 (60.5)
33-36	339	.153 (52)	.132 (44.7)	.139 (47)	.130 (44.2)
37-40	276	.163 (45)	.168 (46.2)	.185 (51)	.170 (47.0)
41-44	176	.273 (48)	.211 (37.0)	.222 (39)	.219 (38.6)
45-48	117	.214 (25)	.261 (30.5)	.248 (29)	.278 (32.5)
49-52	75	.307 (23)	.319 (23.9)	.320 (24)	.345 (25.9)
53-56	28	.357 (10)	.383 (10.7)	.464 (13)	.419 (11.7)
57-60	14	.357 (5)	.451 (6.3)	.357 (5)	.496 (6.9)
	2,442				

* Fraction of fire days with number of fire days in parentheses.

		Seeley Lake District (Seeley Lake ERC's) Frac. fire days (days)		Superior District (Superior ERC's) Frac. fire days (days)	
ERC	Days	Observed	Estimated	Days	Observed	Estimated
0-8	61	.033 (2)	.026 (1.6)	103	.010 (1)	.009 (1.0)
9-16	134	.030 (4)	.042 (5.6)	221	.023 (5)	.021 (4.6)
17-24	334	.063 (21)	.065 (21.7)	481	.035 (17)	.046 (22.1)
25-32	520	.104 (54)	.100 (52.0)	670	.101 (68)	.098 (65.4)
33-36	251	.139 (35)	.137 (34.3)	327	.171 (56)	.166 (54.2)
37-40	142	.169 (24)	.167 (23.7)	251	.275 (69)	.229 (57.6)
41-44	90	.178 (16)	.202 (18.2)	173	.260 (45)	.309 (53.4)
45-48	42	.357 (15)	.242 (10.2)	86	.360 (31)	.401 (34.4)
49-52	28	.179 (5)	.288 (8.1)	14	.500 (7)	.501 (7.0)
53-56	5	.200 (1)	.339 (1.7)	2	1.000 (2)	.601 (.8)
	1,607			2,328		

	Plains District (Plains ERC's) Frac. fire days (days)			Thompson Falls District (Thompson Falls ERC's) Frac. fire days (days)		
ERC	Days	Observed	Estimated	Days	Observed	Estimated
0-8	32	.000 (0)	.013 (.4)	24	.042 (1)	.013 (.3)
9-16	104	.010 (1)	.019 (1.9)	109	.000 (0)	.020 (2.2)
17-24	218	.005 (1)	.027 (5.9)	243	.025 (6)	.031 (7.5)
25-32	476	.034 (16)	.039 (18.5)	369	.049 (18)	.048 (17.6)
33-36	291	.058 (17)	.051 (14.9)	149	.094 (14)	.066 (9.8)
37-40	278	.068 (19)	.061 (17.0)	143	.091 (13)	.081 (11.6)
41-44	261	.084 (22)	.073 (19.1)	144	.104 (15)	.100 (14.3)
45-48	239	.109 (26)	.087 (20.9)	126	.135 (17)	.122 (15.3)
49-52	172	.110 (19)	.104 (17.9)	98	.102 (10)	.148 (14.5)
53-56	111	.090 (10)	.123 (13.6)	64	.141 (9)	.179 (11.4)
57-60	92	.163 (15)	.145 (13.3)	33	.212 (7)	.214 (7.1)
61-64	68	.253 (17)	.170 (11.6)	22	.318 (7)	.255 (5.6)
65-68	42	.238 (10)	.199 (8.4)	11	.455 (5)	.300 (3.3)
69-72	21	.143 (3)	.231 (4.9)	8	.125 (1)	.349 (2.8)
73-76	16	.000 (0)	.267 (4.3)	4	.500 (2)	.402 (1.6)
77-80	15	.267 (4)	.306 (4.6)			
81-85	8	.000 (0)	.348 (2.8)			
	2,444			1,547		

MODELS BASED ON ERC'S FROM THE MISSOULA STATION

	Seeley Lake District (Missoula ERC's) Frac. fire days (days)			Superior District (Missoula ERC's) Frac. fire days (days)	
ERC	Days	Observed	Estimated	Observed	Estimated
0-8	78	.000 (0)*	.010 (.8)*	.000 (0)	.012 (1.0)
9-16	215	.019 (4)	.019 (4.0)	.000 (0)	.024 (5.2)
17-24	418	.031 (13)	.033 (13.7)	.043 (18)	.046 (19.4)
25-32	706	.052 (37)	.058 (40.8)	.082 (58)	.087 (61.5)
33-36	339	.103 (35)	.087 (29.5)	.186 (63)	.137 (46.4)
37-40	276	.120 (33)	.114 (31.3)	.199 (55)	.182 (50.2)
41-44	176	.159 (28)	.147 (25.8)	.261 (46)	.237 (41.8)
45-48	117	.188 (22)	.188 (22.0)	.256 (30)	.304 (35.6)
49-52	75	.240 (18)	.237 (17.8)	.307 (23)	.380 (28.5)
53-56	28	.143 (4)	.294 (8.2)	.429 (12)	.462 (12.9)
57-60	14	.357 (5)	.359 (5.0)	.357 (5)	.546 (7.6)
	2,442				

* Fraction of fire days with number of fire days in parentheses.

		Plains District (Missoula ERC's) Frac. fire days (days)		Thompson Falls District (Missoula ERC's) Frac. fire days (days)	
ERC	Days	Observed	Estimated	Observed	Estimated
0-8	78	.000 (0)	.011 (.9)	.000 (0)	.006 (.4)
9-16	215	.014 (3)	.020 (4.2)	.000 (0)	.011 (2.3)
17-24	418	.026 (11)	.033 (13.8)	.014 (6)	.020 (8.3)
25-32	706	.044 (31)	.055 (39.1)	.040 (28)	.037 (26.3)
33-36	339	.112 (38)	.081 (27.4)	.065 (22)	.059 (20.1)
37-40	276	.134 (37)	.103 (28.5)	.109 (30)	.080 (22.1)
41-44	176	.148 (26)	.131 (23.1)	.114 (20)	.107 (18.9)
45-48	117	.188 (22)	.165 (19.3)	.128 (15)	.142 (16.7)
49-52	75	.133 (10)	.205 (15.4)	.133 (10)	.187 (14.0)
53-56	28	.179 (5)	.253 (7.1)	.179 (5)	.241 (6.7)
57-60	14	.000 (0)	.307 (4.3)	.286 (4)	.305 (4.3)
	2,442				

EXAMPLES OF POISSON MODELS FOR THE SUPERIOR DISTRICT

To illustrate the use of a Poisson model for the number of fires on a day, the logistic model for the Superior District constructed using Superior's ERC's will be used.

$$\hat{P} = \frac{e^{-5.05698 + .10120 \cdot ERC}}{1 + e^{-5.05698 + .10120 \cdot ERC}}$$

ERC	$\hat{P}$	$\hat{\lambda} = -\ln(1-\hat{P})$ = Poisson parameter
20	.046	.047
38	.229	.260

If Z = number of fires on a day, then with

$$\hat{\lambda} = .047 \qquad P(Z=z) = \frac{.047^z \, e^{-.047}}{z!} \qquad z = 0, 1, 2, \ldots$$

and similarly for $\hat{\lambda} = .260$.

$\hat{\lambda}$ = .047 (ERC = 20)

z = number of fires	0	1	2	3	4	5	6	7	≥8	
P(z) = estimated fraction of days	.954	.045	.001							
Observed fraction of days	.965	.025	.002	.004	.000	.002	.000	.000	.002	
Observed number of days	464	12	1	2	0	1	0	0	1	481
Estimated number of days	458.9	21.6	.5	0	0	0	0	0	0	481

$\hat{\lambda}$ = .260 (ERC = 38)

z = number of fires	0	1	2	3	4	5	6	7	≥8	
P(z) = estimated fraction of days	.771	.200	.026	.002						
Observed fraction of days	.725	.187	.052	.012	.004	.008	.004	.004	.008	
Observed number of days	182	47	13	3	1	2	1	1	2	251
Estimated number of days	193.5	50.2	6.5	.5	0	0	0	0	0	251

SUMMARY STATISTICS FOR LOLO NATIONAL FOREST DAYS

For most of the remainder of this report Lolo National Forest is treated as an entity and districts are ignored. The Missoula weather station was chosen as the weather station for these forest days. Before discussing some logistic and Poisson models, a few summary statistics are given.

Lolo National Forest Days

Number of days with a fire	831 (34.0%)
Number of days with a large fire	65 (2.7%)
Total number of days	2,442

No. of fires	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
No of days	1611	393	186	92	48	19	18	15	17	8	9	3	0	3	0	6	2	3

No. of fires	18	19	20	21	22	23	24-36	37	38-48	49	50-57	58	
No. of days	0	0	1	0	3	2	0	1	0	1	0	1	2,442

Correlation of Fire Indexes for the Missoula Station

	ERC	IC	BI	SC
IC	.7302			
BI	.7669	.8352		
SC	.3219	.5370	.7566	
RHINDEX	.6811	.7837	.6447	.3056

The sample size for each correlation coefficient is 2,442.

LOGISTIC MODELS FOR THE PROBABILITY OF A FIRE DAY AS A FUNCTION OF ONE OR TWO FIRE INDEXES FOR LOLO FOREST DAYS

MODELS WITH ONE FIRE INDEX

Each model has the following form:

Probability of a fire = $\hat{P}(x) = e^{c + dx}/(1 + e^{c + dx})$.

Burning Index		c	d
ERC	Energy Release Component	-3.48770	.09110
IC	Ignition Component	-1.57866	.04720
SC	Spread Component	-1.00176	.04952
BI	Burning Index	-2.27212	.04836
RHINDEX	100-Relative Humidity	-2.87284	.03636

MODELS WITH TWO INDEXES

Since the logistic model with ERC was the best model with one index, two variable logistic models were constructed with ERC and one other index in them. Each model has the form:

Probability of a fire day = $\hat{P}(x_1, x_2) = e^{c+dx_1+ ex_2}/(1+e^{c+dx_1+ex_2})$.

Indexes	c	d	e
ERC, IC	-3.64958	.09170	.00358
ERC, SC	-3.70446	.09928	-.01714
ERC, BI	-3.70108	.10110	-.00520
ERC, RHINDEX	-3.69948	.09582	-.00016

The next several tables illustrate how well these models work by comparing observed results with estimates made using these models.

MODELS FOR LOLO NATIONAL FOREST DAYS

ERC	Days	Fraction of fire days (days) Observed	Estimated (Using $\hat{P}$(ERC))
0-8	78	.038 (3)*	.042 (3.3)*
9-16	215	.070 (15)	.084 (18.0)
17-24	418	.165 (69)	.159 (66.5)
25-32	706	.263 (186)	.282 (198.8)
33-36	339	.440 (149)	.404 (136.8)
37-40	276	.522 (144)	.494 (136.2)
41-44	176	.614 (108)	.584 (102.8)
45-48	117	.641 (75)	.669 (78.3)
49-52	75	.667 (50)	.744 (55.8)
53-56	28	.750 (21)	.807 (22.6)
57-60	14	.786 (11)	.858 (12.0)
	2,442		

* Fraction of fire days with number of days in parentheses.

IC	Days	Fraction of fire days (days) Observed	Estimated (Using $\hat{P}$(IC))
0	165	.145 (24)	.171 (28.2)
1-4	209	.110 (23)	.185 (38.6)
5-8	279	.183 (51)	.215 (60.0)
9-12	299	.244 (73)	.249 (74.3)
13-16	287	.275 (79)	.285 (81.9)
17-20	278	.342 (95)	.325 (90.5)
21-24	235	.447 (105)	.368 (86.5)
25-28	205	.459 (94)	.413 (84.7)
29-32	114	.579 (66)	.460 (52.4)
33-36	97	.546 (53)	.507 (49.1)
37-40	77	.545 (42)	.554 (42.6)
41-50	93	.613 (57)	.633 (58.9)
51-60	53	.642 (34)	.735 (34.9)
61-70	19	.789 (15)	.816 (15.5)
71-80	16	.625 (10)	.876 (14.0)
81-90	13	.615 (8)	.919 (12.0)
91-100	3	.667 (2)	.948 (2.8)
	2,442		

SC	Days	Fraction of fire days (days) Observed	Estimated (Using $\hat{P}$(SC))
0-2	168	.113 (19)	.278 (46.8)
3	259	.158 (41)	.299 (77.4)
4	439	.271 (119)	.309 (135.8)
5	430	.353 (152)	.320 (137.6)
6	289	.433 (125)	.331 (95.6)
7	204	.436 (89)	.342 (69.7)
8	119	.378 (45)	.353 (42.0)
9-12	305	.416 (127)	.376 (114.7)
13-20	187	.487 (91)	.448 (83.8)
21-40	37	.622 (23)	.619 (22.9)
41-60	0	.000 (0)	.814 (0.0)
61-100	3	.000 (0)	.951 (2.9)
101-	2	.000 (0)	.988 (2.0)
	2,442		

BI	Days	Fraction of fire days (days) Observed	Estimated (Using $\hat{P}$(BI))
1-8	132	.121 (16)	.111 (14.7)
9-16	97	.062 (6)	.156 (15.1)
17-24	342	.143 (49)	.213 (73.0)
25-28	346	.231 (80)	.266 (92.0)
29-32	443	.323 (143)	.305 (135.3)
33-36	286	.451 (129)	.348 (99.5)
37-40	241	.469 (113)	.393 (94.7)
41-44	164	.457 (75)	.440 (72.2)
45-48	92	.533 (49)	.488 (44.9)
49-52	93	.505 (47)	.536 (49.9)
53-60	105	.562 (59)	.607 (63.8)
61-70	57	.667 (38)	.705 (40.2)
71-80	22	.773 (17)	.795 (17.5)
81-90	17	.529 (9)	.863 (14.7)
91-100	4	.250 (1)	.928 (3.7)
	2,441		

RHINDEX	Days	Fraction of fire days (days) Observed	Estimated (Using $\hat{P}$(RHINDEX))
0-10	20	.050 (1)	.064 (1.3)
11-20	87	.184 (16)	.089 (7.7)
21-30	111	.180 (20)	.123 (13.7)
31-36	92	.163 (15)	.158 (14.5)
37-42	154	.162 (25)	.189 (29.1)
43-48	150	.240 (36)	.225 (33.7)
49-52	134	.201 (27)	.258 (34.6)
53-56	163	.245 (40)	.287 (46.8)
57-60	172	.326 (56)	.318 (54.7)
61-64	225	.280 (63)	.350 (78.8)
65-68	235	.332 (78)	.384 (90.2)
69-72	276	.424 (117)	.419 (115.6)
73-76	248	.480 (119)	.455 (112.7)
77-80	169	.562 (95)	.491 (82.9)
81-90	200	.595 (119)	.554 (110.8)
91-100	6	.667 (4)	.641 (3.8)
	2,442		

A COMPARISON OF THREE LOGISTIC MODELS FOR THE PROBABILITY OF A FIRE DAY IN THE LOLO NATIONAL FOREST

In each model discussed in this section, $\hat{P}$ is the probability of a fire day, ERC is the Energy Release Component, and IC is the Ignition Component.

MODEL 1

$\hat{P}$ is a function of ERC alone. Data were supplied to the logistic program as grouped data, i.e., summarized by intervals of ERC values as in the tables of the preceding section. This model was used in the preceding section.

$$\hat{Y}_1 = -3.48770 + .09110 \cdot ERC \quad , \quad \hat{P}_1 = e^{\hat{Y}_1} / (1 + e^{\hat{Y}_1}) \quad .$$

MODEL 2

$\hat{P}$ is a function of ERC alone. Data were supplied to the logistic program as individual cases (days), i.e., no grouping.

$$\hat{Y}_2 = -3.70376 + .09564 \cdot ERC \quad , \quad \hat{P}_2 = e^{\hat{Y}_2} / (1 + e^{\hat{Y}_2}) \quad .$$

MODEL 3

$\hat{P}$ is a function of ERC and IC. Data were supplied to the logistic program as individual cases (days).

$$\hat{Y}_3 = -3.64958 + .09170 \cdot ERC + .00358 \cdot IC \qquad \hat{P}_3 = e^{\hat{Y}_3} / (1 + e^{\hat{Y}_3}).$$

A comparison of the estimated values for P using $\hat{P}_1$, $\hat{P}_2$, and $\hat{P}_3$ for a selection of values for ERC and IC is given in the following table. Of course, $\hat{P}_1$ and $\hat{P}_2$ do not depend on IC. It is seen that there are only slight differences between $\hat{P}_1$ and $\hat{P}_2$, indicating that method of supplying data to the logistic program has little effect on the model. In model 3, changing the value of IC has little effect on the value of $\hat{P}_3$, indicating that ERC is the dominant variable. This was also true for the other models using ERC and a second variable.

COMPARISON OF THREE MODELS* FOR ESTIMATING P

ERC	IC	$\hat{P}_1$	$\hat{P}_2$	$\hat{P}_3$
10	0	.071	.060	.061
25	0	.230	.212	.205
50	0	.744	.746	.718
75	0	.966	.970	.962
10	10	.071	.060	.063
25	10	.230	.212	.211
50	10	.744	.746	.725
75	10	.966	.970	.963
10	25	.071	.060	.066
25	25	.230	.212	.220
50	25	.744	.746	.736
75	25	.966	.970	.965
50	50	.744	.746	.753

* See the material preceding the table for a definition of the models.

EXAMPLES OF POISSON MODELS FOR LOLO NATIONAL FOREST DAYS

To illustrate the use of a Poisson model for the number of fires on a day, the logistic model with index variables ERC is used.

$$\hat{Y} = -3.48770 + .09110 \cdot \text{ERC} \qquad \hat{P} = e^{\hat{Y}} / (1 + e^{\hat{Y}}) \quad .$$

ERC	$\hat{P}$	$\hat{\lambda} = -\ln(1-\hat{P})$ = Poisson parameter
20	.159	.173
38	.494	.681

If Z = number of fires on a day, then with $\hat{\lambda} = .173$

$$P(Z = z) = \frac{.173^z \quad e^{-.173}}{z!} \qquad z = 0, 1, 2, \ldots$$

$\hat{\lambda}$ = .173 (ERC = 20)

z = number of fires	0	1	2	3	4	5	6	7	≥8	
P(z) = estimated fraction of days	.841	.146	.013	.001						
Observed fraction of days	.835	.120	.024	.014	.002	.000	.000	.002	.002	
Observed number of days	349	50	10	6	1	0	0	1	1	418
Estimated number of days	351.5	61.0	5.4	.4	0	0	0	0	0	418

$\hat{\lambda}$ = .494 (ERC = 38)

z = number of fires	0	1	2	3	4	5	6	7	≥8	
P(z) = estimated fraction of days	.506	.345	.117	.027	.005	.001				
Observed fraction of days	.478	.236	.127	.047	.040	.018	.007	.004	.043	
Observed number of days	132	65	35	13	11	5	2	1	12	276
Estimated number of days	139.7	95.2	32.3	7.5	1.4	.3	0	0	0	276

LOGISTIC MODELS FOR THE PROBABILITY OF A LARGE FIRE DAY FOR LOLO FOREST DAYS

Although the number of large fire days was small (65, 2.7%), it was possible to develop fairly reasonable models for the probability of a large fire day. Each model has the following form:

$\hat{P}$ = probability of a large fire day = $e^{c+dx}/(1 + e^{c+dx})$.

Burning Index	c	d
ERC	-6.07562	.07380
IC	-4.58924	.04072
SC	-3.88544	.03660
BI	-5.26252	.04464
RHINDEX	-3.79737	.06040

The next several tables illustrate how well these models work by comparing observed results with estimated results.

MODELS FOR LOLO NATIONAL FOREST DAYS

		Fraction of large fire days (days)	
ERC	Days	Observed	Estimated (Using $\hat{P}$(ERC))
0-8	78	.000 (0)*	.003 (.2)
9-16	215	.000 (0)	.006 (1.2)
17-24	418	.007 (3)	.010 (4.2)
25-32	706	.020 (14)	.018 (12.6)
33-36	339	.038 (13)	.027 (9.3)
37-40	276	.036 (10)	.037 (10.1)
41-44	176	.045 (8)	.049 (8.5)
45-48	117	.068 (8)	.064 (7.5)
49-52	75	.067 (5)	.084 (6.3)
53-56	28	.071 (2)	.110 (3.1)
57-60	14	.143 (2)	.142 (2.0)
	2,442		

* Fraction of fire days with number of days in parentheses.

IC	Days	Fraction of large fire days (days) Observed	Estimated (Using $\hat{P}$(IC))
0	165	.000 (0)	.010 (1.7)
1-4	209	.000 (0)	.011 (2.3)
5-8	279	.011 (3)	.013 (3.6)
9-12	299	.010 (3)	.015 (4.5)
13-16	287	.007 (2)	.018 (5.1)
17-20	278	.022 (6)	.021 (5.8)
21-24	235	.026 (6)	.024 (5.7)
25-28	205	.049 (10)	.028 (5.8)
29-32	114	.070 (8)	.033 (3.8)
33-36	97	.062 (6)	.039 (3.8)
37-40	77	.078 (6)	.046 (3.5)
41-50	93	.043 (4)	.060 (5.6)
51-60	53	.113 (6)	.087 (4.6)
61-70	19	.105 (2)	.125 (2.4)
71-80	16	.063 (1)	.177 (2.8)
81-90	13	.154 (2)	.244 (3.2)
91-100	3	.000 (0)	.327 (1.0)
	2,442		

SC	Days	Fraction of large fire days (days) Observed	Estimated (Using $\hat{P}$(SC))
0-2	168	.000 (0)	.021 (3.5)
3	259	.000 (0)	.022 (5.8)
4	439	.014 (6)	.023 (10.2)
5	430	.033 (14)	.024 (10.4)
6	289	.024 (7)	.025 (7.2)
7	204	.029 (6)	.026 (5.3)
8	119	.034 (4)	.027 (3.2)
9-12	305	.036 (11)	.029 (8.8)
13-20	187	.064 (12)	.036 (6.7)
21-40	37	.135 (5)	.058 (2.1)
41-60	0	.000 (0)	.114 (0.0)
61-100	3	.000 (0)	.277 (.8)
101-	2	.000 (0)	.535 (1.1)
	2,442		

BI	Days	Fraction of large fire days (days) Observed	Estimated (Using $\hat{P}$(BI))
1-8	132	.000 (0)	.006 (.8)
9-16	97	.000 (0)	.009 (.9)
17-24	342	.003 (1)	.013 (4.3)
25-28	346	.009 (3)	.016 (5.6)
29-32	443	.029 (13)	.019 (8.6)
33-36	286	.035 (10)	.023 (6.6)
37-40	241	.029 (7)	.027 (6.6)
41-44	164	.037 (6)	.032 (5.4)
45-48	92	.033 (3)	.039 (3.6)
49-52	93	.032 (3)	.046 (4.3)
53-60	105	.086 (9)	.059 (6.2)
61-70	57	.105 (6)	.086 (4.9)
71-80	22	.045 (1)	.128 (2.8)
81-90	17	.176 (3)	.187 (3.2)
91-100	4	.000 (0)	.310 (1.2)
	2,441		

RHINDEX	Days	Fraction of large fire days (days) Observed	Estimated (Using $\hat{P}$(RHINDEX))
0-10	20	.000 (0)	.000 (0)
11-20	87	.000 (0)	.001 (.1)
21-30	111	.009 (1)	.002 (.3)
31-36	92	.000 (0)	.004 (.3)
37-42	154	.006 (1)	.005 (.8)
43-48	150	.013 (2)	.008 (1.1)
49-52	134	.007 (1)	.010 (1.4)
53-56	163	.000 (0)	.013 (2.1)
57-60	172	.012 (2)	.016 (2.8)
61-64	225	.022 (5)	.021 (4.7)
65-68	235	.017 (4)	.026 (6.2)
69-72	276	.033 (9)	.033 (9.2)
73-76	248	.065 (16)	.042 (10.4)
77-80	169	.059 (10)	.053 (9.0)
81-90	200	.070 (14)	.079 (15.7)
91-100	6	.000 (0)	.135 (.8)
	2,442		

SCORES FOR EVALUATING PREDICTIONS FOR THE PROBABILITY OF A FIRE DAY USING THE LOGISTIC MODELS

Each logistic model in this report gives an estimate $\hat{P}$ for each day that depends on the values of one or more fire indexes that day where P is the probability of a fire day. In this section a method of evaluating these estimates is discussed.

Each day let E be a random variable with value 1 if a fire occurs and 0 if there is no fire. If $\hat{P}$ is near 1 and E = 1, the prediction is good. Likewise, if $\hat{P}$ is near 0 and E = 0, the prediction is good. Each day a score is calculated as $1 - 2 \cdot (\hat{P} - E)^2$. The quadratic scoring rule for evaluating a model's predictions is given by

$$\text{Average score} = \frac{\sum_i [\, 1 - 2 \cdot (\hat{P}_i - E_i)^2]}{\text{Number of days}}$$

where i runs over all days where predictions are made. The range of possible values for this quadratic score is -1 to +1, with values near 1 being good scores and values near -1 being poor scores. Values of the daily score $1 - 2 \cdot (P - E)^2$ are given in the table that follows for a few values of P and E.

$\hat{P}$	E	$1-2(\hat{P}-E)^2$
1	1	1
0	0	1
.5	1	.5
.5	0	.5
.9	1	.98
.1	0	.98
.9	0	-.62
.1	1	-.62

To compare some of the logistic models constructed in this study, the following procedure was used. Once a logistic model was constructed, it was then used to obtain a $\hat{P}$ value for each of the days used to build the model and the average score was computed. Results of these calculations follow.

SCORES FOR INDIVIDUAL DISTRICTS IN LOLO NATIONAL FOREST

Two logistic models were given earlier for each of the six districts. The first in each case was based on the ERC values for each district, while the second models were based on the ERC values from the Missoula weather station.

District	Index	Average score	Standard deviation of scores	Number of days
Missoula	ERC	.801 (.801)*	.445 (.445)*	2442 (2442)*
Ninemile	ERC	.802 (.802)	.480 (.480)	2442 (2442)
Plains	ERC	.867 (.866)	.420 (.417)	2444 (2441)
Seeley Lake	ERC	.810 (.857)	.460 (.419)	1607 (2442)
Superior	ERC	.797 (.796)	.427 (.433)	2328 (2442)
Thompson Falls	ERC	.858 (.896)	.422 (.380)	1547 (2442)

* The first number in each pair above comes from a logistic model for a district using its own ERC values, while the second number is from a model constructed using Missoula's ERC values.

SCORES FOR MODELS CONSTRUCTED FOR LOLO NATIONAL FOREST DAYS

Several models were presented earlier for Lolo National Forest days. Average scores for each of these models are given in the table below.

Method of data entry	Index(es) in the model	Average score	Standard deviation	Number of days
Grouped	ERC	.627	.365	2442
Grouped	IC	.598	.369	2442
Grouped	SC	.560	.318	2442
Grouped	BI	.593	.353	2442
Grouped	RHINDEX	.587	.342	2442
Individual obs.	ERC, IC	.628	.383	2442
Individual obs.	ERC, SC	.628	.381	2442
Individual obs.	ERC, BI	.628	.381	2442
Individual obs.	ERC, RHINDEX	.628	.382	2442
Individual obs.	ERC	.628	.382	2442

References

1. Martell, D. L., and Otukol, S. 1985. Development of mathematical models for predicting daily people-caused forest fire occurrence in Ontario. Technical Report, School of Forestry, University of Toronto.

2. Winkler, R. L., and Murphy, A. H. 1968. "Good" probability assessors. J. Applied Meteorology. 7: 751-758.

APPENDIX A

Variables Obtained from the National Fire Weather Data Library

Station number, year, month, day, relative humidity, ignition component (IC), spread component (SC), energy release component (ERC), burning index (BI), fire load index (FLI), Julian date, processing option, state of weather, temperature, wind direction, wind speed, minimum temperature, maximum temperature, minimum relative humidity, maximum relative humidity, precipitation duration, precipitation amount, lightning activity level, human-caused risk, 1-hour fuel moisture, 10-hour fuel moisture, 100-hour fuel moisture, 1000-hour fuel moisture, woody moisture, herbaceous moisture, human-caused fire occurrence index, lightning-caused fire occurrence index.

Variables Obtained from the National Fire Occurrence Data Library

District, year, month, day, fire size (0 = no fire this day), statistical cause, general cause, specific cause, agency, Forest Service region, AFFIRMS state code, county, administering forest, watershed number, fire ID, time to discovery, time to initial attack, time to first reinforcement, hours to control, slope, aspect, cover, fuel, area burned, class of people.

APPENDIX B

List of Programs and Data Files

1. WEATHR.DTA — Raw data file of weather/indexes data.
2. FIRE.DTA — Raw data file of fires data.
3. WEATHR.SPS — Sorts weather/indexes data file (WEATHR.DTA) and creates an SPSSX system file WEATHR.SYS.
4. WEATHR.SYS — Weather/indexes SPSSX system file created by WEATHR.SPS.
5. WEATHR1.SPS — Uses WEATHR.SYS and does summary statistics on all weather/indexes variables for checking purposes.
6. FIRE.SPS — Sorts fire data file (FIRE.DTA) and creates system file FIRE.SYS. Assigns station number to each district. Does summary statistics for checking, too.
7. FIRE.SYS — Fires system file created by FIRE.SPS.
8. MERGE.SPS — Attaches weather/index data from WEATHR.SYS to each fire in FIRE.SYS creating a merged system file MERGE.SYS.
9. MERGE.SYS — Merged system file created by MERGE.SPS.

10. MERGE2.SPS — Uses MERGE.SYS to create MERGE2.SYS. A subset of variables is chosen from MERGE.SYS and also only those cases with weather data are kept. This last item is accomplished by dropping cases with no ERC present.

11. MERGE2.SYS — The system file created by MERGE2.SPS.

12. ANAL1.SPS — Uses MERGE2.SYS. Aggregates the file MERGE2.SYS to one case each day in each district creating system file ANAL1.SYS. Variables created: FIRES, LARFIRES, FIREDAY, LFIREDAY.

13. ANAL1.SYS — System file created by ANAL1.SPS.

14. ANAL2.SPS — Uses ANAL1.SYS to get frequency/percentile statistics for most of the variables in this file by district.

15. TEST3.SPS — Uses ANAL1.SYS. Generates raw data needed to use as input to SPSSX's PROBIT to get logistic equations for each district using each district's own ERC values.

16. LOGIS3.SPS — Uses data from TEST3.SPS to run logistic equation for District 3, using its own ERC values.

17. LOGIS4.SPS — Same as LOGIS3.SPS for District 4.

18. LOGIS5.SPS — Same as LOGIS3.SPS for District 5.

19. LOGIS6.SPS — Same as LOGIS3.SPS for District 6.

20. LOGIS7.SPS — Same as LOGIS3.SPS for District 7.

21. LOGIS8.SPS Same as LOGIS3.SPS for District 8.
(Runs of LOGIS3.SPS and LOGIS4.SPS using natural log of ERC rather than ERC were also made.)

22. WEATHRC.SPS Takes only weather/indexes from Station Number 241507 (Missoula) from WEATHR.DTA. Sorts data and creates WEATHRC.SYS.

23. WEATHRC.SYS System file created by WEATHRC.SPS.

24. FIREC.SPS Sorts fires data file FIRE.DTA, assigns same station number (241507) to each district, runs summary statistics for checking and creates FIREC.SYS.

25. FIREC.SYS System file created by FIREC.SPS.

26. MERGEC.SPS Attaches weather/index data from WEATHRC.SYS to each fire in FIREC.SYS creating merged system file MERGEC.SYS. All cases have weather/index information from station 241507.

27. MERGEC.SYS System file created by MERGEC.SPS.

28. MERGE2C.SPS Uses MERGEC.SYS to create MERGE2C.SYS. A subset of variables is chosen and cases with no ERC are dropped. All cases use Station Number 241507.

29. MERGE2C.SYS System file created by MERGE2C.SPS.

30. ANAL1C.SPS — Uses MERGE2C.SYS. Aggregates the file MERGE2C.SYS to one case each day in each district creating system file ANAL1C.SYS. Variables created: FIRES, LARFIRES, FIREDAY, LFIREDAY. Weather/indexes all from Station Number 241507.

31. ANAL1C.SYS — System file created by ANAL1C.SPS.

32. TEST3C.SPS — Uses ANAL1C.SYS. Generates raw data that are used as input to SPSSX's PROBIT to get logistic equations for each district using ERC's from Missoula Station in each district.

33. LOGIS3C.SPS — Uses data from TEST3C.SPS to run logistic equations for District 3 using ERC values from Missoula Station (241507).

34. LOGIS4C.SPS — Same as LOGIS3C.SPS except for District 4.

35. LOGIS5C.SPS — Same as LOGIS3C.SPS except for District 5.

36. LOGIS6C.SPS — Same as LOGIS3C.SPS except for District 6.

37. LOGIS7C.SPS — Same as LOGIS3C.SPS except for District 7.

38. LOGIS8C.SPS — Same as LOGIS3C.SPS except for District 8.

39. ANAL1CA.SPS — Uses MERG2C.SYS. Aggregates the file MERGE2C.SYS to one case each day in Lolo National Forest creating system file ANAL1CA.SYS. Variables created: FIRE, LARFIRES, FIREDAY, LFIREDAY. Lolo Forest days--weather/indexes Station Number 241507.

40. ANAL1CA.SYS — System file created by ANAL1CA.SPS.

41. TEST3CA.SPS Uses ANAL1CA.SYS to generate raw data needed to use as input to SPSSX's PROBIT to get logistic equations for Lolo Forest days using ERC, IC, SC, BI, RHINDEX=100-RELHUMID. Weather/indexes from Missoula Station (241507). (For firedays and large fire days.)

42. LOGISLO.SPS Uses raw data from TEST3CA.SPS to run logistic equations for probability of a fire day and probability of a large fire day for Lolo Forest Days, Missoula Station weather/indexes. Variables (ERC, IC, SC, BI, RHINDEX) are used one at a time, i.e., only one variable in each model.

43. LOGIS9CA.SPS Same as LOGISLO.SPS except only for ln(ERC).

44. TWOVAR1.SPS Uses ANAL1CA.SYS. Does crosstabs of ERC vs. IC.

45. TVARLOG2.SPS Uses ANAL1CA.SYS. Logistic of FIREDAY with ERC using individual cases rather than grouped data. Missoula Station weather.

46. TVARLOG1.SPS Uses ANAL1CA.SYS. Does logistic of FIREDAY with two variables--ERC, IC. Missoula weather/indexes. Individual cases used.

47. TVARLOG3.SPS Same as TVARLOGI.SPS except ERC, SC used.

48. TVARLOG4.SPS Same as TVARLOG1.SPS except ERC, BI used.

49. TVARLOG5.SPS Same as TVARLOG1.SPS except ERC, RHINDEX=100-RELHUMID used.

50. ERCCORR.SPS Uses ANAL1.SYS. Gives correlations between ERC values at the six districts in Lolo National Forest.

51. INDXCORR.SPS Uses ANAL1CA.SYS. Computes correlations of all pairs from ERC, IC, SC, BI, FLI, RELHUMID from Missoula Station (241507).

52. SCORES1.SPS Uses ANAL1CA.SYS. Computes quadratic scores for logistic models for Lolo Forest days. Each model has only one variable in it. ERC, BI, IC, SC, or RHINDEX.

53. SCORES2.SPS Uses ANAL1CA.SYS. Computes quadratic scores for logistic models for Lolo Forest days. Each model has two variables in it. (ERC, BI), (ERC, SC), (ERC, IC), or (ERC, RHINDEX).

54. SCORES3.SPS Uses ANAL1.SYS. Computes quadratic scores for logistic models for ERC logistic models in each district with each district's model using that district's ERC values.

55. SCORES4.SPS Uses ANAL1C.SYS. Computes quadratic scores for logistic models for ERC logistic models in each district with each district's model using Missoula Station's ERC values.

APPENDIX C

Items Worthy of Further Study

This study was exploratory in nature. Some very interesting and promising results were obtained. A number of additional items stand out as possible directions to go in further research. Some of the more important of these are listed briefly in this appendix.

1. Other fuel models should be tried in the Lolo National Forest models. For fuel models that are less appropriate than the one used in this study, the logistic models and quadratic scores should not be as good as those obtained in this study.

2. For Lolo National Forest, separate models could be built for human-caused and lightning-caused fires. The majority of the fires in Lolo National Forest are caused by lightning.

3. Further study should be done on using quadratic scores to judge logistic models. This seems to be a very promising method for judging these models.

4. The approach developed in this study should be tried on another Forest that is substantially different than Lolo National Forest, e.g., in Mississippi. In Mississippi, there are many more fires, most fires are human-caused, and different seasons are likely to be an important factor.

5. A journal article should be written for publication.

6. There have been some very recent articles on making inferences for logistic models. Some of these results may be useful since logistic models are a key element in the approach of this report.

7. The procedures of this report that were carried out using SPSSX should be converted to procedures using SAS on the Data General when this is possible. The logistic procedures in SAS are better and more sophisticated than those in SPSSX.

8. The use of more than two variables in the logistic models should be investigated. This study looked at a few models with two variables and none with more than two variables.

9. If the approach used in this study continues to give good results when applied to a wider range of applications, work should begin on ways to implement the procedures in practice.